

Contents

Ton van den Beld Introduction.....	1
Part I. Moral Responsibility and Ontology: Setting the Stage	
Peter van Inwagen Moral Responsibility and Ontology.....	11
R. Jay Wallace Moral Responsibility and the Practical Point of View	25
Part II. Responsibility and Personal Identity	
Keith Graham Collective Responsibility.....	49
Marc Slors Personal Identity and Responsibility for Past Actions.....	63
Eric Wiland Personal Identity and Quasi-Responsibility.....	77
Part III. Incompatibilist Arguments	
Kasper Lippert-Rasmussen Does Moral Responsibility Presuppose Alternate Possibilities?.....	89
David Mackie Fischer on Alternative Possibilities and Responsibility.....	103
Maureen Sie Freedom and Blameworthiness.....	113
Christopher Mark Grau Moral Responsibility and Wolf's Ability.....	129
René van Woudenberg Moral Responsibility and Agent Causation	143
Part IV. Compatibilist Positions	
Ish Haji On the Value of Ultimate Responsibility	155
James Lenman Contracting Responsibility.....	171

Kevin Magill
Blaming, Understanding and Justification 183

Paul Russell
Compatibilist-Fatalism 199

Part V. Some Related Issues

Aaron Ben-Ze'ev
Emotions, Responsibility and Morality..... 219

Angela M. Smith
Identification and Responsibility 233

Michael S. McKenna
Toward a Speaker Meaning Theory of Moral Responsibility..... 247

Stefaan E. Cuypers
Alfred Mele's Voluntaristic Conception of Autonomy..... 259

Biographical notes 271

Index..... 273

Freedom and Blameworthiness

Maureen Sie

Introduction

In this contribution I will argue that the intelligibility of the category of actions that I call 'exemplary blameworthy' depends upon the existence of robust alternative possibilities. This necessity of alternative possibilities is no problem for those who accept a conditional or hypothetical analysis of 'could have done otherwise,' although these views have notorious problems of their own. It is a problem, though, for those who want to accept Frankfurt's famous counter-examples to the principle of alternative possibilities or for those who want to establish compatibilism along the lines suggested by these counter-examples. It is this kind of compatibilism that is the focus of this contribution.

The basic problem fueling the discussion on the (in)compatibility of what is called 'deep responsibility' and determinism, is that determinism renders all our actions, as Peter van Inwagen has put it, equally the consequences of the laws of nature and of what happened in the remote past.¹ But it is not up to us what went on before we were born, nor is it up to us what the laws of nature are. Therefore, the consequences of these things, including our present acts, are not up to us.² Determinism then - the theory that all our mental events, including our choices and decisions, as well as our actions, *are effects of certain things* and therefore have to happen or are *necessitated*, and cannot be owed to *origination*³ - poses a problem. If everything is equally the inevitable product of past events together with the laws of nature, the question is why robots, young children and the psychologically less capable are not as deeply - if at all - responsible for their actions as we, so-called normal human beings are. If determinism is true, after all, we all act and behave *equally* according to a specific set of unchosen determinants.

¹ 'Deep responsibility' is distinguished from 'superficial responsibility.' The latter refers to our 'merely causal' responsibility for events; i.e. the kind of responsibility we share with natural phenomena, animals and artefacts.

² P. Van Inwagen, *An Essay on Free Will* (Oxford: Clarendon Press, 1983), 16.

³ T. Honderich, *How Free are You? The Determinism Problem* (Oxford: Oxford UP, 1993), 138.

According to such diverse thinkers as Harry Frankfurt, Susan Wolf and R. Jay Wallace the solution must be sought in certain general abilities that - unlike us - robots, young children and the mentally incapacitated lack or are unable to exercise. If we could enable the robot to determine the content of its will⁴, to control its movements reflectively⁵, or to act in accordance with Reason⁶, then the robot would become as deeply responsible for its actions as we - normal adult human beings - normally are.

Frankfurt's hierarchical account is one of the most influential compatibilistic ones of this century.⁷ It connects our responsibility for a certain action *A* exclusively with our 'decisive identification' to do *A*.⁸ If we did *A* because we really wanted to do *A* then we are responsible for *A*, regardless of the availability of so-called robust alternative possibilities.⁹ Hence, the truth of determinism - which, according to most philosophers, excludes the existence of alternative possibilities - is irrelevant to our responsibility for certain of our actions. This is what is meant to be established by the famous Frankfurt-style counterexamples to the Principle of Alternate Possibilities - the principle that states that someone is only morally responsible for *A* if she could have done otherwise than *A*. If someone really wants to do *A* and does *A* because of this, she is morally responsible for it whether or not such counterfactual interveners as the omnipotent scientist called 'Black' or more lay forms of overdetermination, ensure that the agent can do nothing other than *A*.

Unfortunately, this account has several familiar problems. One of these is, to state it very briefly, that it is unclear *why* the ability to form and act in accordance with a so-called higher-order volition makes us deeply responsible for the resulting action. Another is that it is unclear how hierarchical authorization accounts can accommodate the existence of mitigating and excusing circumstances; circumstances that lessen the blame that is someone's due, *independent* of the fact that the agent acted on the basis of a decisive identification. The accounts provided by both Wolf and Wallace are partly meant to overcome these shortcomings.

⁴ H. G. Frankfurt, 'Freedom of the Will and the Concept of a Person', *The importance of what we care about* (Cambridge: Cambridge UP, 1988), 11-25.

⁵ R. J. Wallace, *Responsibility and the Moral Sentiments* (Cambridge: Harvard University Press, 1994).

⁶ S. Wolf, *Freedom within Reason* (Oxford: Oxford UP, 1990).

⁷ The hierarchical model was, at the same time, also developed by Gerald Dworkin, but is primarily known through the work of Frankfurt. G. Dworkin, *The theory and practice of autonomy* (Cambridge: Cambridge UP, 1988).

⁸ H. G. Frankfurt, 'Identification and Wholeheartedness', *The importance of what we care about*, 159-176.

⁹ H. G. Frankfurt, 'Alternate Possibilities and Moral Responsibility', *The importance of what we care about*, 1-10; H. G. Frankfurt, 'Coercion and Moral Responsibility', op. cit., 26-46.

Our ability to 'control our behaviour reflectively' or to 'act in accordance with reason' enables us to 'do the right thing for the right reasons,' as Wolf puts it, or, as Wallace puts it, to 'avoid culpable choices.' Hence, since culpable choices and wrong actions are what we hold one another responsible for, the possession of the abilities to avoid wrong actions and culpable choices immediately explains our deep responsibility for these actions. Unlike robots and very young children, we full-blown responsible human beings possess the ability to do what we should do, hence, deserve to be blamed for our failure to do so.

If, on the other hand, certain circumstances or conditions interfere with or undermine our ability to act in accordance with reason or to control our behaviour reflectively, we can be partly or fully exonerated from the blame we would deserve in normal circumstances.¹⁰ Under these conditions we are no longer able to do what we are generally able to do, and therefore it would be unreasonable to (fully) blame us for what we did.

Although I find these kinds of compatibilist accounts very compelling and believe that they can be defended adequately in the end, the aim of this contribution is to point out a fundamental difficulty that all accounts of deep responsibility in terms of the possession of certain general abilities share. This difficulty concerns the intelligibility of the claim that an agent possesses the responsibility-relevant abilities, even though her action is an exemplary blameworthy one and therefore evidence to the contrary. This difficulty, I will argue, can only be solved by the assumption of robust alternative possibilities.

That exemplary blameworthy actions are 'evidence to the contrary' - that they are indistinguishable from actions of agents who do not possess the responsibility-relevant abilities - is what I will argue for in the first section of this contribution. I will define an exemplary blameworthy action in formal terms and will show how we can disregard many possible disagreements about the content of this definition. In the second section, I will elaborate on the fundamental difficulty with the category of blame-worthy actions, and will argue that this leads to the necessity of the assumption that alternative possibilities exist. In the third section, I will discuss two possible ways to avoid this difficulty, and conclude that these two ways do not work.

¹⁰ Wolf's view and Wallace's view differ greatly in this as in other respects. Wallace restricts the exonerating or exempting circumstances to those that invalidate the powers of reflective self-control itself, whereas Wolf contends that the mitigating or exonerating circumstances include those in which we 'lack the *opportunity* to exercise the abilities relevant to responsibility.' S. Wolf, op. cit., 101-102; R. J. Wallace, op. cit., 187-188, 223-225.

I Exemplary Blameworthy Actions

If we are deeply responsible human beings in virtue of our reflective or evaluational abilities an 'exemplary blameworthy action' should be defined as 'a wrong action of a human being who possesses the relevant abilities to act in a morally appropriate way and who does not have any excuse not to.' In this section, I will argue that this definition should be divided into the following four claims. Action *A* of person *P* is blameworthy iff: 1) *P* is a normal human being, 2) *A* is a true action of *P*, 3) *A* is something that no normal human being should do, and 4) *P* did *A* under normal circumstances.

I.1 Normality. The first part of our definition, which states that 'P is a normal human being,' guarantees that the agent of an exemplary blameworthy action has no exceptional characteristics - whatever these are - on the grounds of which she can be exempted from bearing responsibility for the blameworthy action.

One possible source of controversy with regard to responsibility in general and my definition of blameworthiness in particular, concerns the precise abilities, capacities, or characteristics that make an agent a responsible subject. However, it is not important to the argument of this contribution how we exactly define the class of responsible human beings, how we determine who belongs to it, or which actions such beings are able to perform. Hence, we can put aside the possible controversy about the abilities that are relevant to responsibility (hereafter: 'RR abilities') by using the ambiguous phrase 'normal human beings' to indicate the possession of RR abilities, regardless of which abilities these are.

I am aware of the statistical connotations of the term 'a normal human being' but I believe that, first of all, the notion shares and illuminates many of the ambiguities that characterize the notion of 'responsible human being.' Secondly, it enables the problematic, almost incoherent nature of a blameworthy action to be explicated.¹¹

What is important for the argument of this contribution is that the classification 'normal' equals the assumption that one is able to behave and act in certain specified ways, whereas the classification 'abnormal' or 'deviant' - which can be the result of, for instance, age (e.g. very young children) or mental health - equals the assumption that one is incapable of behaving and acting in certain specified ways.

Although I will speak about 'normal' and 'deviant' human beings in general terms, people are often normal or deviant with respect to certain

¹¹ It also, thirdly, suggests a way to understand the kind of Compatibilism discussed in this contribution, as a third option distinct from both traditional Incompatibilism and traditional Compatibilism. I elaborate on this in the extended version of this contribution.

(classes of) actions and behaviour only, because they possess or lack a very specific ability.¹² For instance, someone with Tourette's syndrome lacks the ability to behave quietly and decently all the time, and is therefore 'abnormal' - but only in regard to behaving and acting in accordance with the norms of decency.¹³ In this sense, the condition that an exemplary blameworthy action must have been performed by a normal human being is not controversial. If your general constitution makes you unable to perform certain acts or unable not to perform certain acts, then you are not to blame for failing to perform or for performing them, respectively.¹⁴

I.2. Ownership. The second part of our definition - i.e. that 'A is a true action of P' - guarantees that the exemplary blameworthy action has no special features - whatever these are - on the grounds of which A cannot be used to evaluate the agent.

This feature is very difficult to spell out. On the one hand we want to exclude such actions as those that are carried out under the influence of, for instance, hypnosis or manipulation and also those actions that are sheer accidents or that are unforeseen side-effects of another intentional action. On the other hand we don't want to exclude such actions as, for instance, 'accidents caused by a reckless person' or 'grave harm inflicted while under the influence of alcohol.' The problem is that we lack a clear and distinct definition of the things on the basis of which we can and should be evaluated; worse, we even lack agreement about which things *should be* the basis.

Fortunately, we need not solve this controversy for we can avoid it by stipulating that the class of actions with which we are concerned is that of the actions that are 'potential vehicles of self-disclosure,' a phrase used by Paul Benson (although Benson reserves it explicitly for 'completely free actions' only).¹⁵ What we exclude by defining this class thus, are those actions that do not tell us anything, not even potentially, about the agent. Hence, the emphasis in 'potential vehicles of self-disclosure' is on 'poten-

¹² Cf. Benson who suggests that the grammar of freedom is perhaps 'relational.' P. Benson, 'Freedom and Value', *The Journal of Philosophy* 84 (September 1987), 486.

¹³ The case of Tourette's syndrome is especially worrisome because the 'inability' concerned is difficult to define, and relative. People who suffer from the syndrome are often able to constrain themselves and to control their 'uncontrollable' impulses, but only with extreme effort. Some people are even able to 'integrate' the syndrome to a high degree in their personality. Hence, although it is not literally impossible for them to behave quietly and inconspicuously, it would be unreasonable to expect it of them because it would demand an inhuman effort. Hence, our expectations of one another take into account not only our individual abilities but also the effort it costs us to exercise them.

¹⁴ Cf. B. Gert, T. J. Duggan, 'Free Will as the Ability to Will', *Noûs* 13 (1979), 197-217.

¹⁵ P. Benson, op. cit., 465-486

tial.' It excludes actions on the basis of which we cannot be evaluated, but includes all actions on the basis of which we *might* be evaluated - which is not to say that all the actions on the basis of which we might be evaluated are actions on the basis of which we should be evaluated. Most actions we perform are unsuitable to be and worthless as *actual* vehicles of self-disclosure.

If I trip on the carpet and spill my coffee, my spilling of the coffee is not a vehicle of self-disclosure if it *really* was an accident that I could not avoid; although it *might* be if I am a very careless person and my tripping was due to 'not taking care.' Which actions should and which should not count as potential vehicles of self-disclosure is not important for the argument of this contribution, as long as we agree that there exists a distinction between actions that allow us to evaluate the agent and those that do not. With regard to the latter class of actions, it is not controversial that one should not blame agents for actions that do not tell us anything about them.

I.3. Normative Expectations. The third part of our definition - which states that 'A is something that no normal human being should do' - ensures that the blameworthy action is one that does not correspond to our 'legitimate expectations' - whichever these are - of a normal human being.

These so-called expectations must be understood as 'general expectations,' in other words, those that are grounded in our belief that a certain individual is a person or an adult human being like ourselves. This includes those expectations we have of someone as a human being, as well as the more specific expectations we have of someone as, for instance, a parent or - more specifically - the expectations we have of someone as a colleague, a neighbour or a 'friend of mine.' It excludes those expectations we have or form on the grounds of a certain individual's specific characteristics; for instance, the expectation that Aunt Mary will make a joke when asked the time, because she always makes a joke when asked the time. This latter expectation is merely predictive, whereas the expectations that, for instance, 'people answer questions,' 'neighbours help us out' and 'friends stick together' are not primarily predictive, but also normative. Not only do we expect people to answer a question, but we believe that the correct way to respond to a question is to give an answer. The expectations that are breached in the case of a blameworthy action are always normative expectations, i.e. expectations concerning what 'should be done' in certain circumstances.

Not every breach of a normative expectation, though, is an opportunity to evaluate the individual agent; sometimes it is an opportunity to evaluate the content of the expectation that is breached. Some of our normative expectations could well prove to be inappropriate, unjustified or inadequate if, for instance, we discover that they were informed only by some unexamined local habit of the group to which we belong and the contingent circumstances in which we find ourselves. Some of our normative expect-

tations, then, are not legitimate or are not legitimate without some further specification or refinement. If such an expectation is breached we should not blame the individual agent, but adapt or reformulate our expectations. Hence, in the case of an exemplary blameworthy action the normative expectations that are breached must be *legitimate* normative expectations, expectations that Wallace has dubbed 'the moral demands that we accept.'¹⁶

Although the question which moral demands are and which are not legitimate might be a source of controversy, this is not important to the argument of this contribution.¹⁷ It might also be a source of controversy whether, and if so in what sense, there really are norms and values on the basis of which we have legitimate normative expectations. Here, we can follow Wolf who has rightly argued that the minimal kind of value realism necessitated by the existence of such expectations and demands must be shared by anyone who wants to discuss deep responsibility at all. The true value-sceptic cannot but abandon all talk about praise and blame, for she does not acknowledge the existence of things for which we could be blamed or praised.¹⁸ Hence, if we believe that a discussion on the existence of deep responsibility is intelligible at all, we must also assume that certain values exist and that these values make certain normative expectations legitimate and others not. Once we accept this, it can hardly be controversial that an exemplary blameworthy action must be an action that breaches a normative expectation that we believe to be legitimate.

I.4. Mitigating Circumstances. The last part of our definition - i.e. that '*P* did *A* in normal circumstances' - establishes that the exemplary blameworthy action is not due to exonerating or mitigating personal circumstances - whatever these are - such that the agent can be partly or fully excused from blame for the action.

Of all the actions of a normal human being that potentially disclose something about her and that do not correspond to our legitimate normative expectations - of all the *wrong actions* - the really *blameworthy actions* are those for which there are no exonerating or mitigating personal circumstances such that the agent can be partly or fully excused from blame for it. It is at this point that 'merely' wrong actions can be distinguished from truly blameworthy ones in a more than arbitrary way.

¹⁶ R. J. Wallace, *op. cit.*, 41.

¹⁷ Elsewhere I argue that the general assumption that we are deeply responsible human beings can be justified on the basis of the necessity of a certain modesty with respect to the legitimacy of our normative expectations, see Maureen Sie, *Responsibility, Blameworthy Actions and Normative Disagreements* (Utrecht University, Ph.D. thesis, 1999), chapter 4.

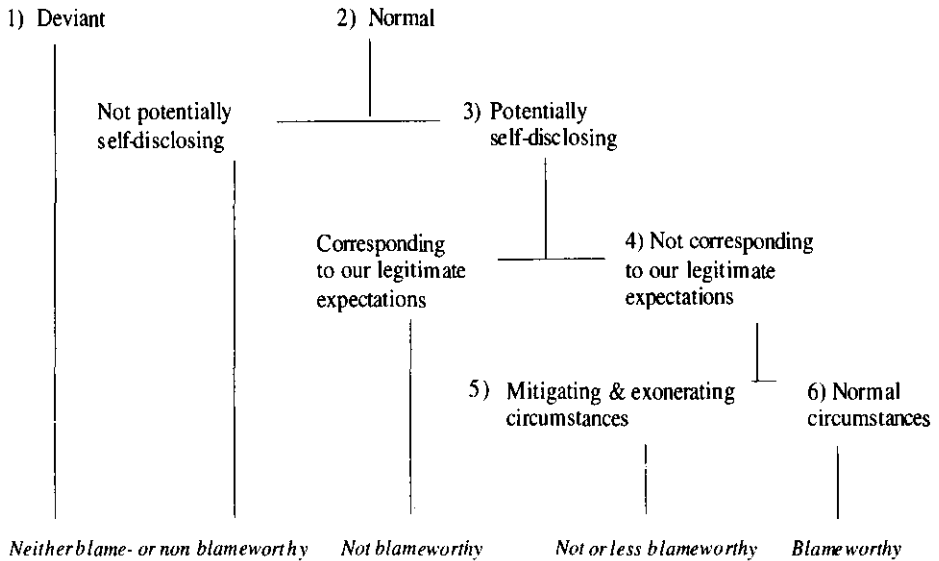
¹⁸ S. Wolf, *op. cit.*, chapter 6.

In comparison with the previous category, a wrong action carried out under mitigating, excusing or exonerating circumstances won't incline us to reconsider the legitimacy of the content of our normative expectations. For example, imagine a woman who drinks a few glasses of wine and is then asked by a neighbour to watch a sick child. When the temperature of the child rises slightly, the woman decides to take the child to the hospital, 'just in case.' She gets into her car despite the few glasses of wine, and drives the child to the hospital. She believes that one can never be too sure, that a few glasses of wine do not make her unfit to drive a car, that her neighbour will be very grateful, and so on.

Nothing in this situation invites us to reconsider the demand that 'one should under no circumstances drive under the influence of alcohol.' The woman's reasons for driving despite the few glasses of wine are not *good* reasons: the child is not really that sick, and 'over-protectiveness' - especially if it leads to potentially dangerous situations - is not a good character trait. However, if we are told that this woman lost her only child just a few years ago, this will mitigate our blame. Explanations that cite exceptional personal circumstances - such as the loss of one's only child - lessen the blame because, and in as far as, they make the wrong action, somehow, understandable without making it any less wrong, though.

If the woman keeps on acting in ways we disapprove of on occasions that involve children, we will adapt our classification of her as a normal human being with regard to 'actions that involve children.' Consequently, we will adapt our normative expectations of her as an individual agent. More often than not, though, a wrong action that - somehow - seems to be explainable by exceptional circumstances will not lead to such a reclassification, but only to a milder moral evaluation. Hence, we accept a class of actions that are wrong but not (fully) blameworthy.

The following is the model with which we start and that is compatible with a lot of different accounts of responsibility.



Within the category 'normal,' on the right side of the model, there is a subclass of agents that act in a potentially self-disclosing way, but contrary to our legitimate normative expectations (4). Within that class we have a subclass of 'wrongdoers' (5) and a subclass of agents who deserve to be fully blamed for their wrong action, i.e. the class of 'blameworthy agents' (6).

The difficulty now that this contribution aims at, is how to account for the distinction between the sixth category of agents who act in a blameworthy manner and the first category of deviant individuals, *against the background of the distinction between wrongdoing and blameworthiness* as it is made above. In order for an action to be not only exemplary wrong but also exemplary blameworthy it must deviate *inexplicably* from what we expect of a normal human being. After all, if it can be explained by exceptional circumstances it is no longer inexplicable, but no longer exemplary blameworthy either. But if someone behaves in such an inexplicable deviant manner how can we then maintain that she is a normal human being, i.e. that she does not belong to the first category of deviant individuals? This question discloses the problematic - or, as I will argue, paradoxical - nature of the concept of an exemplary blameworthy action, which can only be remedied when we assume the existence of robust alternative possibilities. Let me explain this in more detail.

II Freedom as alternative possibilities

Following the model step-by-step we see that the deviancy of the agent that is inferred on the grounds of a specific deviant action cannot be explained away by a specification or redescription of the action on the grounds of which we judge the agent to be deviant.

That it is the individual who is judged, claimed or experienced to be at once 'deviant' *and* 'normal' is stipulated by the fact that her action belongs to the class of potentially self-disclosing actions (3). This means that the action can be used to evaluate the agent. If this condition were not satisfied, the action would fall into the class on the left side of the model, i.e. the class of 'not potentially' self-disclosing actions.

That the individual who acts wrongly is, in at least one sense, deviant is stipulated by the fact that her behaviour does not correspond to our legitimate normative expectations of her (4). If a certain redescription of the action were possible, we would reclassify it as an action that falls into a class on the left side of the model, i.e. an action that *corresponds* to our legitimate expectations. Suppose, for instance, that we turn the television on and the first shot we see is a 'killing,' but then the next shot communicates that the killer is actually a charmingly handsome hero and the setting is a cold, cruel war. In this case we would immediately reclassify the action 'killing' as one that corresponds to our legitimate expectations, by redescribing it as 'a killing in time of war.'

That the individual must be deviant in a robust sense - that is, really different from that which we expect from a normal human being - is stipulated because the class of actions that do not correspond to our normative expectations but that we can nevertheless come to understand fully or at least partially constitute a separate class, again on the left side of the model: that of the 'not (or less) blameworthy actions.' These actions are wrong but the agent who performed them is not (fully) to blame. As said, the woman from the previous section who drives the child to hospital after she had drunk a few glasses of wine acts wrongly, but we can understand what motivated her to act in a way we disapprove of.

Hence, in order to arrive at the class of exemplary blameworthy actions conceptually, we must accept the following premises:

- 1) there are normal human beings;
- 2) they sometimes act in ways that enable us to evaluate them;
- 3) some of these potentially self-disclosing actions are wrong;
- 4) we cannot understand these wrong actions on the grounds of any exceptional circumstances that obtain.

Hence:

- 5) there are normal human beings who act in ways that really deviate from how we legitimately expect a normal human being to act.

Premise 5, however, implies that there are occasions on which an individual agent 'could have done otherwise than she in fact did' in exactly the same circumstances, for someone who did wrong must have been able to act in a way that corresponds to the way she was enabled to act by the sum total of her abilities relative to the sum total of the circumstances, in order to be not only a wrongdoer, but also an individual who acts in an exemplary blameworthy manner.

If no-one is ever able to act otherwise than they in fact do someone's deviant actions are always evidence of the fact that some exceptional personal circumstance obtained, or of the fact that she is a deviant human being, i.e., that she is not as able as we thought her to be, but is suffering from certain mental, physical or volitional deficiencies that explain her deviant action. If the latter is the case, we should adapt our classification of the individual accordingly.

In defense of the kind of compatibilism that is under scrutiny here, one could object that it is absurd to problematize someone's status as a 'normal human being' or as a 'deviant' one because it is clear and unmistakable who among us possesses the responsibility relevant abilities and who does not. But this is only true, if at all, in cases where the *malfunctioning* of an individual is very obvious or when we are able to identify the lack of certain abilities independent of the malfunctioning itself, for instance by genetic, brain or other neurophysiological defects. What we are looking for in the case of an exemplary blameworthy action, though, is a *well-functioning* human being whose individual action is indistinguishable from that of a malfunctioning human being.

The problem is that a seemingly normal human being might suffer from various forms of inability that are unobservable, for instance, because of their complete integration into her personality. If this is the case, her prima facie blameworthy action tells us something about her, but what it tells us is not that she is a normal human being - i.e. someone who possesses RR abilities - who acts reprehensibly, but that she is a deviant human being - i.e. someone who does not possess RR abilities in sufficient degree.

In as far as deep responsibility is concerned, the distinction between normal and deviant human beings is problematic. Of course we are able to make a rough distinction between 'normal' and 'deviant' without difficulty; but we are also able to make a rough distinction between 'responsible' and 'non-responsible' human beings. The controversial question with regard to deep responsibility and its existence is whether these distinctions that play such an important role in our daily lives are ultimately intelligible if determinism is true - i.e. whether the condition that accounts for these distinctions is compatible with determinism. In this respect, the worrisome cases are the borderline ones in which it is not so clear whether or not someone's behaviour is the result of the sum total of her abilities and the circumstances she is in.

It is, as the sceptic will point out, a slippery slope that leads us from 'some doubt' about particular cases to 'general doubt,' in this case, about the existence of a firm distinction between normal and deviant human beings. And the sceptic is right: there are so many types of mental disease and so many syndromes and they cover such a wide scope of human actions and behaviour - from our ability to spell correctly (dyslexia) to our ability to act and behave in a morally appropriate way (Phineas P. Gage) - that it is hard to make a firm and clear distinction between normal and deviant human beings.

Especially in the case of blameworthy agents, the difficulty is that we have no way to define what it means for someone to possess certain abilities (relative to the circumstances), unless we base the definition on the behaviour those abilities enable her to exhibit in these circumstances. In the case of a wrong action, though, the behaviour that we believe an individual to be capable of is typically absent. Therefore, when someone acts contrary to our legitimate expectations - the circumstances of action being normal and non-excusing - the question is when and for what reasons our classification of her as a normal human being was not mistaken. For, as Dennett has put it, 'the grounds for saying that the person is culpable (the evidence that he did wrong, was aware he was doing wrong, and did wrong of his own free will) are in themselves grounds for doubting that it is a person we are dealing with at all.'¹⁹

Hence, if we believe that there are exemplary blameworthy actions as I have defined them, we must conclude that we are able to distinguish someone's status as a normal human being from her (deviant) behaviour. That unproblematic conclusion, however, implies the problematic conclusion that there must exist people who acted contrary to our normative expectations, but were nevertheless able to act consistently with our normative expectations, unconditionally (i.e. under exactly the same circumstances).

Let me discuss two possible ways to avoid the strong conclusion that freedom as alternative possibilities is necessary for the intelligibility of the category of individuals who act in a blameworthy manner.

III Two easy ways out

The necessity of robust alternative possibilities would disappear if we are able to distinguish or define someone's status as a normal human being, apart from the actions that her being so actually enables her to perform. If someone is a normal human being only if she is able to act and behave in

¹⁹ D. C. Dennett, 'Conditions of Personhood', *The Identities of Persons*, A.O. Rorty (ed.) (Berkeley: University of California Press, 1976), 194.

certain ways and if her acting and behaving in certain ways is what defines her as a normal human being, then every deviant action would prove that the individual agent is in exceptional circumstances or lacks the necessary abilities. Hence, one way to look for an account of blameworthiness without the existence of robust alternative possibilities is to look for a way to define someone's status as a normal human being, distinct from the behaviour and actions that she actually performs.

Several objections to the necessity of alternative possibility as I have defined it spring to mind. Due to the limited space available, I will discuss only two of them, and only briefly at that.

The first objection that springs to mind is that someone's status as a normal human being is a historic phenomenon that depends upon the totality of her actions and behaviour. Hence, to say of someone that she acted reprehensibly boils down to saying that someone who acts and behaves as most of us do most of the time (or with regard to most things), acted reprehensibly this time (or with regard to this thing). According to this reply, the difficulty I have pointed out is based upon the misconception that we are ever confronted with the time-slice phenomenon of an individual and her isolated, blameworthy action. In everyday life, so the objection will run, we know which individuals are normal and which are deviant because we always see a picture bigger than that of the single action alone.

Unfortunately, this objection only amounts to a solution to our problem if we can accept that only those who act 'out of character' can act really reprehensibly. This is implausible. A car accident caused by the negligence of a reckless driver who lives a reckless life seems a pretty good rather than a bad example of an exemplary blameworthy action.

Hence, acting normally 'most of the time' is not necessary for one's ability to act in an exemplary blameworthy way. Nor does it seem sufficient. One can act normally most of the time with regard to most actions and still lack certain abilities that other people have, or one may possess these abilities but only to such an extent that certain conditions prevent one from using them. Someone who has the ability to pass an exam can nevertheless flunk it due to the specific circumstances in which the exam was taken; for instance, a certain aspect of the situation made her unable to concentrate sufficiently because of some unacknowledged childhood-syndrome. If this is the case, she is not to blame for failing the exam.

Although the historic view might give us a good epistemic account of *how* we come to judge each other as 'deviant' or 'normal' individuals in daily life, it doesn't give us an account of the conditions of 'normality,' distinct from the 'behaviour that it makes someone capable of' - which is what we need in order to make the category of exemplary blameworthy actions intelligible without a robust notion of freedom as alternative possibilities.

The second objection is that we should distinguish between *motivating* reasons and *justifying* reasons, i.e., perhaps recognizing and understanding the reasons behind an action (as motivating reasons) makes us judge that an action is an action of a normal human being, whereas our rejection of these motivating reasons as justifying reasons makes us judge the action to be a deviant one. In this case, the difficulty that we are examining - the apparent incoherent, even paradoxical, nature of an exemplary blameworthy agent - disappears because the agent's deviation is not contradicting that which gives the agent her status as a normal human being in the first place. Clearly, we recognize the motivating reasons of some actions, although, they run counter to our normative expectations because we do not accept these motivating reasons as justifying reasons. Although we understand the attraction of, for instance, tax evasion, we do not condone people who evade paying tax. Hence, perhaps we hold on to an individual's status as a 'normal human being' because, in a sense, we expect her deviant behaviour from a normal human being, although we do not approve of it: we want every human being to behave a good deal better than that.

If this is true, and I believe that there is much to say in favour of this, the class of exemplary blameworthy actions consists (entirely) of those actions that we partly understand but nevertheless disapprove of. Hence, the person who kills the neighbourhood cats because they kill her pigeons, acts really reprehensibly. It is no wonder that someone becomes very angry if a cat kills her pigeons, but killing the cats is the wrong reaction to it. If, on the other hand, the person kills the cats without an intelligible motivating reason, we would probably reconsider our assumption that she is a normal individual and conclude that she is out of her mind or insane.²⁰ Apparently she does not possess RR abilities, for no-one in her right mind kills neighbourhood cats just for fun.

If we understand the class of blameworthy actions with the help of the distinction between motivating and justifying reasons, it becomes 'smaller' than the traditional class of exemplary blameworthy actions, because it excludes all truly 'moral monster'-like crimes we usually allow and use as examples of blameworthy actions. Although I believe that there are good reasons to accept this restriction it isn't of much help with regard to the problem at hand. The distinction between justifying and motivating reasons only obscures the difficulty under scrutiny by replacing it with the familiar concepts of 'weakness of will' and, a sub-class of this, 'laziness of mind.'

²⁰ There is a difficulty with regard to our moral sentiments towards so-called moral monsters, but I do not think it would be very difficult to provide an error-account in these cases.

If we recognize the reasons for someone's deviant behaviour as motivating reasons and if we recognize these motivating reasons as not being justifying reasons, then again we need to assume the wrongdoer to be capable of recognizing the same, notwithstanding the fact that she did not recognize it (laziness of mind). Or, in the case that she does recognize the motivating reasons as not being justifying reasons, we should assume her to be capable of acting upon this recognition, notwithstanding the fact that she did not act upon it (weakness of will). Hence, the labels 'weakness of will' and 'laziness of mind' are an abbreviation for the complete formula that contains the explanandum of this contribution.

The labels 'laziness of mind' and 'weakness of will' suggest that a certain effort wasn't made although it was perfectly legitimate to expect that certain effort to be made by that person at that moment and with regard to the circumstances she was in. Therefore, the necessity of the assumption of alternative possibilities remains the same; i.e. we assume, and we need to assume, that our perfectly reasonable and legitimate expectations were not fulfilled by someone who could have fulfilled them (unconditionally), for otherwise it would be wiser to adapt our expectations of the individual who acts in a blameworthy manner. Again, the crucial distinction is made between the behaviour that does not fulfill our legitimate expectations and the agent who is perfectly capable of fulfilling those expectations.

Conclusion

Accepting an account of deep responsibility in terms of the possession of some RR abilities does not side-step the difficult issue of the (in)-compatibility of alternative possibilities and determinism. Without the assumption that at least some people are sometimes able to act in a way that corresponds to our normative expectations even if they did not actually do so, we must abandon the idea of a category of blameworthy actions, i.e. we must abandon the idea of a difference that really makes a difference between wrong actions and blameworthy ones.

If someone doesn't live up to our normative expectations, the conclusion that she 'failed to behave in ways she is capable of' - as opposed to the conclusion that 'our expectations of her were unjustified' - depends upon the assumption that she was free or able to act - or to try to act - otherwise than she in fact did, in exactly the same circumstances. If we conceive of the necessity of alternative possibilities in this way, determinism is threatening because it renders the distinction between a 'normal' and a 'deviant' human being equal to the distinction between a 'well-functioning' and a 'malfunctioning' human being - just as it renders the distinction between 'normal' and 'exceptional' circumstances, equal to the distinction

between 'suitable' circumstances for well-functioning and 'unsuitable' circumstances for well functioning.

According to this account of the necessity of alternative possibilities, the crucial question is whether the sum total of our abilities together with the sum total of the circumstances produce our actions - just as the sum total of the currents constituting the flood causes the poorly constructed dam to break.²¹ If this is so, all our actions are 'mere effects' in a causal chain that started long before we were born, and people are only superficially responsible for the consequences of their actions - just as the poorly constructed dam is only superficially responsible for failing to avert the flood.

If, on the other hand, the possession of certain RR abilities enables us to originate events, we must conclude that some events can only be understood by referring to an individual agent who did what she did, chose as she chose, decided as she decided, and did so freely - i.e. she could have done, chosen or decided to do otherwise. Since determinism so obviously excludes origination and we - in our daily practices of responsibility - seem to accept the existence of blameworthy actions, the burden of proof on the compatibilist shoulders remains heavy.

As long as we do not understand *how* the existence of causally sufficient conditions for each and every event allows for the idea of origination by a single individual²² or as long as we cannot explain the distinction between a wrongdoer and a blameworthy agent regardless of this lack of origination,²³ our daily practices of responsibility will be in conflict with determinism, if not incompatible with it. Whether the causes of 'well functioning' or 'malfunctioning' are conceptualized at the level of elementary particles, at that of neurological events and brain states, or at that of psychological and mental phenomena does not matter to the conclusion.²⁴

²¹ R. Chisholm, 'Human Freedom and the Self', *The Lindley Lectures* (Department of Philosophy, University of Kansas, 1964), 25.

²² Primarily directed at Wolf, *op. cit.*, who accepts the necessity of alternate possibilities with regard to the category of blameworthy actions but argues that it is perfectly reasonable to treat it as compatible with determinism until 'otherwise is proven,' because psychological freedom is not necessarily incompatible with physiological determinism, and psychological determinism is not very likely to be true. Cf. criticism of C. Grau, 'Moral Responsibility and S. Wolf's Ability', this volume; M. Ravizza, J. M. Fischer, 'Responsibility, Freedom, and Reason', *Ethics* 102 (Jan. 1992), 385-88.

²³ Primarily directed against R. J. Wallace, *op. cit.*

²⁴ I thank Jan Bransen, Bert van den Brink, Christopher Grau, Marc Slors and Susan Wolf for commenting on and discussing earlier versions of this contribution. I also thank Ton van den Beld for his helpful corrections and comments.